This session examines the validation of highway and transit assignment. *For highway assignment, topics include: Equilibrium assignment definition, volume/delay functions, and multi-class assignment. For transit assignment, topics include transit path building and associated settings, multi-path transit assignment. We introduce the concepts of individual simulation of auto and transit trips.*

Questions and Answers during Session 6

Q: Will the new Validation Report have a functional form for "maximum allowable deviation" (or whatever it's called)?
A: Can you elaborate on what you mean by functional form?
Q: An algebraic function we can put into a spreadsheet to compare a link and identify if it's "in" or "out"
A: I see. I will need to check with Tom about the contents of the manual. I will get back to you on this. Thanks for the question!
Post-session follow-up: It has not yet been decided whether this type of curve will appear in the new manual. If it does, probably both a graph and a formula will be presented.

Q: Where do you get these standards from for future reference?
A: Typically, the standards are based on what is reasonable given all the uncertainties in model data and procedures. The standards are derived more based on consensus than on any exact science.

Q: How is observed VMT derived--count x link length? If so, I see a problem because link length is an arbitrary value and counts are not on every link.
A: This is a very good point. Tom is addressing this point right now. Let us know if you want further clarification.
Post-session follow-up: This could be an issue as not all links are the same length, and longer links are weighted more heavily in VMT estimates. However, the networks are designed, as much as possible, to have links where the volumes do not change along the length of the link. There is indeed more travel on longer links than shorter ones. One way of looking at this is that when you sum a measure, you should not be counting the same thing more than once. Point traffic counts should not be summed because multiple counts can include the same vehicles or trips. VMT, on the other hand, can be summed as the VMT on one link is included only on that link. Point volumes can be summed as long as the probability of double counting is low, as is the case with screenlines. It is important to remember, as said during the presentation, that this is a personal opinion, not a policy or standard. It is presented for your information; you should decide what would be best for your model.

Q: What do you recommend in resort areas with lots of 2nd homes and hotels when doing VMT checks? Thx.
A: This kind of travel should be addressed using a specialized trip purpose called “visitor or tourist.” If data are available on such trips, then they should be incorporated into the modeling process. If this is done, the VMT will automatically reflect such travel. In terms of checking the VMT, HPMS is usually the best source regardless of the type of travel.
Post-session follow-up: Another important thing to consider is that travel models usually represent a specific season, often an average condition such as spring or fall. The peak for recreational travel may be different in specific regions, or parts of regions, and traffic data used for validation should reflect what the model represents.

Q: A model could assign too many trips on one arterial street and very few on other arterial opposite to the counts. But the total or RMSE would be okay. Thus the model is unusable.
A: If this is the case, we need to examine the characteristics of the two arterials in question. That is, why are two arterials that are so close, getting different levels of assignment? Usually, such results are more due to errors in network coding. Of course, there could be issues with the model, but we must also check the network.

Q: That is the Greater Lansing, Michigan, Tri-County TAZ map.
A: Tom must have borrowed it from Lansing then.
Post-session follow-up: I thought it looked familiar, as I was in Lansing the previous week. It actually is an older map that appeared in the 1997 validation manual. I did seem to mention Michigan a lot in the session!

Q: Is there a publication date for the new Validation Manual?
A: I will get back to you on this. I am not aware of the exact dates.
Post-session follow-up: The exact date has not yet been determine. We hope it will be toward the end of this year.

Q: RMSE should be divided by n-1
A: Thanks for this. You are right. In reality though, this may not make a big difference, because the number of links is very high.
Post-session follow-up: If you look up a general definition of %RMSE (not necessarily in a travel modeling context), you will see “n” in the denominator, since it represents the average mean error. However, there are clearly n-1 degrees of freedom. However, as Yasasvi pointed out, the use of n vs. n-1 makes little difference practically. The important thing is to be consistent in all your calculations.

Q: Please comment on the GEH statistic as a better measure than %RMSE
A: The GEH statistic accounts for the magnitude as well as the percentage error. Let me give you an example: if the observed count is 100, and the model predicts 90, the model is off by 10%. If the count is 1000 and the model predicts 900, the model is still off by 10% but in terms of magnitude, it is off by 100 trips instead of 10. This is why the GEH statistic normalizes the squared difference with the average of the observed and predicted volumes.

Q: Is "total screenline volumes" the sum of volumes across all links crossing the screenline?
A: Yes, this is the sum of the volumes criss-crossing the screenline. However, a better check is to look at RMSEs because if we just look at the sums, we may have very high volumes on some and very low on some others, and yet appear to be doing well.

Q: Shouldn't we be looking at difference plots, and difference plots by facility type?
A: Absolutely. This is a very good visual way of checking the volume differences. We could also look at % differences on the map. Thanks for mentioning this.

Q: Are there separate standards for peak hour and peak period models?
A: Usually, the standards are based on facility types and not on time of day.

Q: In addition to individual link speed checks, there is also value in checking zone-to-zone travel times.
A: Thanks for pointing this out.

Q: Could we also use GEH statistic?
A: Sure. GEH statistic is a good way of accounting for magnitude of the difference in addition to the percentage difference. I will touch upon this again at the end, because we had one other participant raise the GEH issue. Thanks.